Receptor Flexibility in de Novo Ligand Design and Docking

Ian L. Alberts,* Nikolay P. Todorov, and Philip M. Dean

De Novo Pharmaceuticals, Compass House, Vision Park, Histon, Cambridge CB4 9ZR, U.K.

Received March 3, 2005

One of the major problems in computational drug design is incorporation of the intrinsic flexibility of protein binding sites. This is particularly crucial in ligand binding events, when induced fit can lead to protein structure rearrangements. As a consequence of the huge conformational space available to protein structures, receptor flexibility is rarely considered in ligand design procedures. In this work, we present an algorithm for integrating protein binding-site flexibility into de novo ligand design and docking processes. The approach allows dynamic rearrangement of amino acid side chains during the docking and design simulations. The impact of protein conformational flexibility is investigated in the docking of highly active inhibitors in the binding sites of acetylcholinesterase and human collagenase (matrix metalloproteinase-1) and in the design of ligands in the S1′ pocket of MMP-1. The results of corresponding simulations for both rigid and flexible binding sites are compared in order to gauge the influence of receptor flexibility in drug discovery protocols.

Introduction

The primary aim of any drug discovery effort is to design novel compounds that are potent toward therapeutically important protein targets. This usually follows either a ligand-based approach if structural information concerning the protein target is not available, but a set of active molecules are known, or a structure-based approach if the three-dimensional structure of the protein target is utilized in the ligand design process. Many rational drug design studies that have proved successful have exploited the holo form of the protein complexed to active molecules.[1] In these protocols, new ligands are often designed by altering the known inhibitors to increase the binding affinity or enhance the specificity of the ligands toward the therapeutic target. The generated ligands typically have similar binding modes to the original inhibitor, although they may represent novel chemotypes. This can lead to several disadvantages. A residue mutation that reduces the affinity of the known inhibitor may also reduce the affinity of the designed ligands. The new potential drug candidates may also have problems concerning toxicity and bioavailability. Furthermore, the above approach precludes the identification of alternative binding modes that may influence properties other than affinity, such as selectivity. The generation of diverse ligands in terms of binding mode as well as structure may be able to circumvent these limitations.

The importance of protein structural flexibility for drug discovery has been scrutinized in an excellent review by Teague.[2] Teague highlights ion channels, nuclear hormone receptors, transporters, and allosteric modulatory sites as classes that undergo conformational changes that are important for design purposes. For example, nonpeptidic piperidine-based inhibitors of renin bind to a receptor conformation, involving concerted rotation of three side chains (Tyr75, Leu73, and Trp39) and movement of the active site flap, not previously known for the binding of peptidic ligands.[3] Conformational rearrangements in proteins can lead to the binding of structurally diverse ligands,[4,5] and, thus, the incorporation of protein structural flexibility into the ligand generation procedure should enhance the diversity of the designed molecules. Protein conformational changes induced upon ligand binding can range from the local rotation of a few side chains to whole domain rearrangements.[6,7]

Many computational approaches to determine the ligand binding mode in a protein binding site, so-called molecular docking techniques, have been developed, and, in general, they are reasonably accurate and efficient.[8−12] Protein conformational flexibility is an important aspect of ligand binding;[13] however, as a consequence of the exponential growth of the number of possible receptor conformations with the number of degrees of freedom, limitations are imposed on the range of conformations sampled during docking. The simplest constraint is to allow no flexibility and utilize a static protein snapshot. This remains the most common recipe in molecular docking although, recently, docking studies have allowed receptor flexibility to some extent.

In general, two main schemes have been used to deal with receptor flexibility in the area of ligand docking. First, an ensemble of predefined receptor conformations can be exploited. Proteins exist in an equilibrated ensemble of structures, and ligand binding involves an equilibrium shift toward the protein−ligand complex structure. Receptor structural models can be obtained from many sources; multiple crystal or NMR structures, molecular dynamics or energy minimizations,[14] and homology models.[15,16] In these approaches, the binding energy of the ligand is assessed against all models in a "cross-docking" protocol.[17−19] In a related approach, the multiple receptor models are averaged, typically by combining interaction grids, and the single averaged structure is used in docking simulations.[20−22] In general, the receptor ensemble approach is found to improve

* Corresponding author. Phone: +44 1223 238054. Fax: +44 1223 238088. E-mail: ian.alberts@denovopharma.com.

10.1021/jm050196j CCC: $30.25
Published on Web 09/22/2005

binding mode predictions and virtual screening enrichment factors.

Second, the receptor conformation has been allowed to change dynamically during the docking simulation. This has been limited to specific degrees of freedom on the protein, such as amino acid residue side chains while keeping the protein backbone fixed, either by identifiying the optimal side-chain torsional angles during the docking procedure[23,24] or using a rotamer library to represent the preferred side-chain orientations.[25−28] In an alternative approach invoking a flexible pharmacophore based scheme, the receptor is represented by grid points onto which pharmacophoric properties are mapped and optimized.[29]

Protein side-chain flexibility in molecular docking has been investigated for several receptors,[30,31] and allowing the rearrangement of several side chains was found to improve the number and diversity of "hits" in a virtual screening protocol.[28] Despite the strong evidence for the impact of receptor flexibility in drug discovery,[32−34] most computational de novo drug design approaches nevertheless treat the protein as a static system, often from a high quality X-ray structure when attempting to design novel, active ligands within the site.[35−38]

Side chains typically exhibit greater mobility than backbones of protein structures, particularly side chains in surface regions,[39,40] and it is clear that incorporation of side-chain flexibility represents an important first step toward treatments of fully flexible protein structures. In this paper we propose a method to incorporate protein flexibility in terms of side-chain rearrangements for both de novo ligand design and docking.

The algorithm utilized in the flexible protein investigations detailed here has the capability of representing side-chain conformers by rotamers or by side-chain χ torsional angles. The rotameric representation was pioneered for side-chain packing studies,[41] and many rotamer libraries have subsequently been developed, including backbone-dependent[42,43] and backbone-independent[44,45] libraries. A number of studies have been reported, which exploit rotamer libraries for homology modeling,[46,47] protein design,[48,49] and ligand docking.[25,26,28] The space of receptor conformations arising from all possible rotameric combinations is often too large to sample exhaustively, and, thus, optimization and clustering techniques have been invoked, including Monte Carlo approaches,[48,50] simulated annealing,[51] dead-end-elimination[52] and its refinements,[53,54] and branch-and-bound schemes.[49]

In protein structure prediction or design, the main aim is often to identify the single lowest energy conformation. In ligand binding explorations, however, it is important to sample low energy structures, since proteins possess conformational entropy. For this reason, the algorithm presented here exploits a simulated annealing protocol to sample receptor side-chain conformational space during the evolving ligand docking or design process. Our approach uses the previously documented and validated drug design program SkelGen[38,55,56] and extends it to include receptor side-chain flexibility, which can be accessed dynamically during the molecular docking or structure generation procedure.

Table 1. Protein Structures Used for Ligand Docking and Generation

receptor	PDB code	resolution/Å	ligand	ref
AChE	1eve	2.5	E2020 (donepezil)	57
AChE	1vot	2.5	huperzine A	58
AChE	1dx4	2.7	tacrine derivative	59
AChE	1qo9	2.7	n/a	59
MMP-1	966c	1.9	RS-104966	66
MMP-1	2tcl	2.2	RO 31-4724	74

The method is used to explore the effect of receptor flexibility in ligand docking and design in the binding site of acetylcholinesterase (AChE) and in the S1′ pocket of human collagenase (matrix metalloproteinase-1 (MMP-1)). Acetylcholinesterase is a serine hydrolase that is involved in a number of therapeutic areas. For example, reversible AChE inhibitors are used to treat myasthenia gravis, an autoimmune disorder characterized by debilitating muscle weakness, and neurodegenerative Alzheimer's disease.[57] Irreversible or competitive AChE inhibitors are highly toxic and include Sarin, Soman, and Tabun. Side-chain rearrangements have been observed between the apo form of AChE and its holo forms and between different ligand-bound forms.[58,59] Two sets of ligand docking experiments are conducted for AChE. The first set involves rotating a single side-chain Phe330, which is known to adopt alternative conformations upon binding of different ligands to AChE. Second, the rearrangement of a large number of side chains is explored for the docking of a tacrine analogue to its native AChE receptor and to the apo form of AChE.

Human collagenase, MMP-1, is a member of the family of enzymes involved in degradation of extracellular matrix proteins. As these are the most abundant proteins in the body, MMPs play a key role in both normal and diseased conditions. MMPs are known to be involved in cancer and pathogenic states, such as arthiritis and arteriosclerosis.[60,61] Several classes of potent MMP inhibitors have been developed, such as carboxylic acid derivatives, hydroxamates, biphenyl and tetracycline analogues.[62] Examination of the known crystal structures of MMP-1−inhibitor complexes reveals that the S1′ pocket undergoes significant conformational changes in multiple side chains to allow the binding of different ligands.[63−65] Indeed, ligand binding selectivity has been ascribed to the structural changes in the S1′ pocket.[66] Two sets of experiments were conducted in the S1′ pocket of MMP-1. First, docking simulations of the known synthetic inhibitor RS-104966 into a flexible S1′ cavity structure were conducted. Second, structures were designed in the flexible S1′ pocket, to a common hydroxamate scaffold that was anchored in the site. We compare docking and structure generation using a static crystal structure receptor and a flexible receptor in all cases.

Methods

Receptor Model Preparation. The AChE and MMP-1 receptor models used in the ligand docking and generation simulations are shown in Table 1. The set for AChE comprises three ligand-bound complexes as well as the native apo form. The two MMP-1 structures are both holo forms. Hydrogen atoms were added to the structures using the molecular modeling package InsightII (Accelrys Inc.), and their atomic positions were minimized within Discover3 to an energy gradient of 0.001 kcal mol^{-1} Å^{-1} using the conjugate gradients

algorithm and the CFF force field,[67] with the non-hydrogen atoms fixed in their crystal structure positions. All ligands, non-Zn metal atoms, and water molecules were removed from our dataset of structures. The corresponding sets of AChE and MMP-1 structures were superposed by alignment of the C^{α} position of the appropriate backbones.

Side-Chain Flexibility. The flexible receptor method either uses combinations of rotamers to represent receptor side-chain conformations or samples side-chain χ torsion angle space, without recourse to rotamers. Both modes of exploring side-chain flexibility were investigated in this work. Many rotamer libraries have been developed for the rotameric representation of side chains, and, in this work, we chose two of the most recent and extensively validated libraries: the May 2002 release of the backbone-dependent rotamer library of Dunbrack and co-workers[42,43] as well as the backbone-independent rotamer library of Xiang and Honig.[44] Both rotamer libraries are constructed from a statistical analysis of a corpus of 844 protein structures. The backbone-dependent rotamer library supplies a register of side-chain χ angles grouped in bins of ± 5 degrees in backbone ϕ, ψ angles. In contrast, the backbone-independent rotamer library supplies a list of side-chain χ angles without recourse to the local backbone ϕ, ψ angles. The backbone dependent library, therefore, provides a relatively small set of rotamers for each flexible residue according to the local backbone orientation, whereas with the backbone-independent library, all rotamers in the library are available for each flexible residue. This allows the influence of backbone torsion angles on side-chain conformations to be explored. For the backbone-dependent rotamer library, to prevent repetitions in the rotamer sets used, only rotamers with associated ϕ, ψ angles within 5° of the original crystal structure were chosen from the library. Furthermore, for both rotamer libraries, only those rotamers that did not clash with any backbone atoms or with any nonflexible side-chain atoms were retained. The van der Waals radii were assigned from the Bondi set,[68] and a heavy atom van der Waals overlap of 0.7 Å was allowed. This value for the van der Waals overlap is consistent with that used in previous SkelGen applications.[28,38] The numbers of possible rotamers for the flexible residues of the different receptors are shown in Table 2.

Table 2. Number of Possible Rotamers for AChE and MMP-1 Crystal Structures from the Backbone-Dependent (bb_{dep}) and Backbone-Independent (bb_{ind}) Rotamer Libraries

AChE

	1vot		1eve	
residue	bb_{dep}	bb_{ind}	bb_{dep}	bb_{ind}
Trp83	4	40	4	42
Trp279	7	47	6	55
Phe330	2	24	3	30

AChE

	1qo9		1dx4	
residue	bb_{dep}	bb_{ind}	bb_{dep}	bb_{ind}
Tyr71	3	67	3	72
Trp83	5	119	5	131
Tyr324	4	50	4	54
Phe330	3	51	2	53
Tyr370	4	67	3	68
Phe371	2	38	3	52
Tyr374	3	71	3	64
Trp472	3	25	3	36
His480	6	132	5	100

MMP-1

	2tcl		966c	
residue	bb_{dep}	bb_{ind}	bb_{dep}	bb_{ind}
Leu81	8	85	6	85
Arg114	21	195	25	171
Val115	4	15	4	15
Ser139	4	25	4	21
Tyr140	6	64	6	86
Phe142	6	63	7	76

Ligand Structure Generation. The de novo drug design program SkelGen was used for the ligand docking and structure generation simulations in this work. The SkelGen algorithm as well as practical applications have been described in detail previously,[38,55,56] and only a summary of the process is given here. The modifications required in order to handle receptor flexibility are highlighted. The input to SkelGen corresponds to the three-dimensional coordinates of the receptor, the coordinates of a rectangular box that defines the boundary of the binding-site pocket, and a set of molecular fragments for the ligand construction process. The algorithm utilizes a stochastic procedure to minimize an objective function that measures the "fit" of putative molecular structures within the binding-site box. An initial structure is built by linking together several fragments in a stepwise manner. The evolving structure is modified by a set of possible transitions, including rigid-body displacements of the whole structure, bond rotations and fragment additions, removals and replacements. The type of transition is randomly chosen at each step, and the value of the objective function is ascertained. The altered structure is accepted or rejected on the basis of the Metropolis condition, such that the probability of acceptance of the transition $p = \exp(\Delta F/T)$, where ΔF corresponds to the change in objective function value before and after the transition and T is analogous to the temperature, which is slowly lowered during the simulation. Thus, the algorithm is guided by the objective function, until a structure is encountered that satisfies all set criteria. The length of each simulated annealing run is characterized by the number of transitions at each temperature, the Markov chain length, and the number of Markov chains as the temperature is gradually lowered. The output generated from a single run is a single ligand, and, as a result of the stochastic nature of the algorithm, different solutions are obtained from different random starts.

SkelGen uses ScreenScore[69] to estimate receptor−ligand binding energy. This is a knowledge-based scoring term that has been optimized to maximize the enrichment in virtual screening.[69] ScreenScore combines aspects from several different scoring functions and has proved to be useful at discriminating between active and inactive compounds.

Modifications for Handling Receptor Flexibility. The conformational space represented by side-chain χ torsion angles or rotamer combinations is sampled through the integration of an additional transition type for side-chain rearrangements in the simulated annealing protocol. Initially a set of χ torsional angles (or a rotamer) is randomly chosen for each of the flexible residues, and, if the new side-chain rotation transition type is picked during the simulation, then one of the flexible residues is chosen at random and a new set of χ torsional angles (or a rotamer) is randomly designated for that residue. The transition is scored and either accepted or rejected on the basis of the Metropolis condition in the same manner as for the other transition types. In this work, we use ScreenScore values to score side-chain conformations, including van der Waals clashes, by measuring the interactions between the flexible side chain and the rest of the protein. The flexible side-chain score for the original crystal structure is determined at the start of the simulation and set as the reference side-chain value. When a side-chain rearrangement transition takes place, a ScreenScore value for the new side-chain conformation is determined and the reference value is subtracted from this. If the difference is less than zero, i.e., the new conformation scores better than the original crystal, then no penalty is added to the overall objective function. On the other hand, a difference greater than zero means that the new conformation is unfavorable compared to the original crystal and the weighted difference is added as a penalty to the objective function. This ensures the formation of low energy conformations compared to the original crystal.

In this study, SkelGen was utilized in both structure generation and molecular docking modes. For structure gen-

Table 3. Binding Energies and Binding Mode for Docking of Donepezil into Its Cocrystal Receptor 1eve and the Non-Native Receptor 1vot

receptor	min energy docking soln/kJ/mol	best rmsd/Å (rank)
rigid 1eve	−38.7	1.0 (1)
rigid 1vot	−35.2	4.5 (1)
flexible 1eve	−39.3	1.0 (1)
flexible 1vot	−39.6	1.3 (1)

eration, the simulation proceeds as discussed above. The predefined molecular template set consists of 1678 fragments derived from fragmentation of compounds in the World Drug Index (WDI) by retrosynthetic analysis and fragmentation patterns of known synthetically feasible molecules.[70] In one set of experiments, molecular structures were generated that were attached to a common hydroxamate scaffold that was anchored in the site. Pharmacophoric constraints were also imposed on the design process to direct the structure generation to key regions of the binding site. Any atoms of the putative ligand should match the pharmacophore restraints; otherwise, penalty terms will be added to the objective function corresponding to the sums of the distances of the nearest atoms to the appropriate pharmacophore sphere.

In SkelGen molecular docking, the ligand to be docked corresponds to the sole fragment used and is individually subject to a SkelGen simulated annealing run. Transitions for fragment addition, deletion, and substitution are, therefore, excluded, and only transitions for ligand conformation changes and side-chain moves are preserved. In all our experiments, the simulated annealing protocol is exploited to dynamically minimize the ScreenScore estimates of the receptor−ligand binding energy, while ensuring that the simulation constraints are satisfied, including removal of all bumps involving ligand and flexible receptor, formation of low energy receptor conformations compared to the original crystal and satisfaction of pharmacophore constraints.

Simulated Annealing Parameters. All de novo ligand structure generation and docking runs for both flexible and rigid receptors used the same set of simulated annealing parameters; 30 Markov chains, each of length 2000, starting temperature value 3.0, temperature scaling factor 0.9. For each ligand design experiment, 500 structures were generated and, similarly, 500 poses were generated for each docking experiment. In terms of computational resources, the flexible receptor simulations are about 5 times slower than the corresponding rigid receptor simulations for the same number of Markov chains, Markov chain length, and annealing cooling schedule. This means that the dynamic protein flexibility method considered in this work scales favorably in comparison to approaches involving receptor conformational ensembles with many members (>5), since each member of the ensemble is basically treated as a single rigid entity.

Results

Acetylcholinesterase−Flexible Ligand Docking. (i) Flexibility of a Single Side Chain. Significant side-chain conformational mobility of the so-called "gateway" residue Phe330 has been observed in crystal structures of AChE with different bound ligands.[57,58] In this experiment, the highly potent commercial drug donepezil was docked into its cocrystal receptor, 1eve, from which the ligand was extracted ("self" docking[71]) as well as a non-native structure 1vot, with the inclusion of side-chain rearrangements of Phe330. Reference donepezil docking runs were also performed into the rigid protein structures. Binding energy and rmsd data for these simulations are presented in Table 3.

Docking of donepezil into the rigid cocrystal form recovered the known crystal structure binding mode of the ligand to within 1.0 Å. However, as depicted in Figure 1, a solution closer than 4.5 Å to the crystal was not found in corresponding docking runs into the non-native structure, as a consequence of clashes with side-chain atoms of Phe330.

Docking into the flexible forms of both receptors by allowing random side-chain rotations of Phe330 yielded the correct ligand binding mode as the top-ranking solution, to within 1.0 Å of the crystal, with the correct side-chain conformation of Phe330 as depicted in Figure 1. Random torsion angles are initially assigned to the flexible residue, and, thus, reproduction of the correct orientation, even for the native receptor, represents a key validation step of the flexible receptor approach. For the non-native state 1vot, the binding site rearranges by rotation of Phe330 to provide sufficient space for accommodation of donepezil. The top-ranking solutions for docking of donepezil into the flexible native and non-native sites, −39.6 and −39.3 kJ/mol, respectively, were lower than that for docking into the rigid native receptor, −35.2 kJ/mol. As a result of the stochastic nature of the algorithm, alternative structures were generated in different binding modes, but these were of higher energy.

The outcome of the docking simulations using rotamers is very dependent on the rotamer library adopted. Using the November 2002 backbone-dependent rotamer library of Dunbrack and co-workers,[43] receptor flexible docking was unable to reproduce the correct ligand binding mode (no solution was within 4 Å of the crystal) or Phe330 side-chain conformation for either the native or the non-native receptors. This is a consequence of the lack of a rotamer with appropriate χ angles (associated with the corresponding backbone ϕ, ψ angles) for Phe330 that matches the crystal structure. In contrast, the backbone-independent rotamer library of Xiang and Honig[44] contains several rotamers with suitable χ angles, and utilization of this rotamer library successfully gave the correct ligand binding mode and Phe330 conformation as the top-ranked solution. It is clear that care must be taken in the choice of rotamer library. Alternatively, the random χ angle (nonrotamer) mode, which eliminates rotamer library dependencies, may be more appropriate.

(ii) Flexibility of a Large Number of Side Chains. Comparison of the structures of the AChE complex with a tacrine derivative, 9-(3-phenylmethylamino)-1,2,3,4-tetrahydroacridine, and the apo form (Figure 2) exemplifies some mobility in several residues,[2] notably the side chains of Tyr71, Tyr370, and Trp83. The other side chains highlighted show less prominent conformational changes. Docking simulations of the tacrine analogue, a highly potent AChE inhibitor, into its cocrystal structure and into the apo form of AChE were conducted. Side-chain flexibility was allowed for Tyr71, Tyr370, and Trp83 as well as for six other residues which show smaller conformational changes: Tyr324, Phe330, Phe371, Tyr374, Trp472, His480. This allows us to explore whether the approach can recover the orientations of mobile as well as more structurally invariant side chains. Binding energy and rmsd data for these runs are presented in Table 4.

Docking simulations to the rigid receptor, 1dx4, from which the ligand was extracted, were able to recover

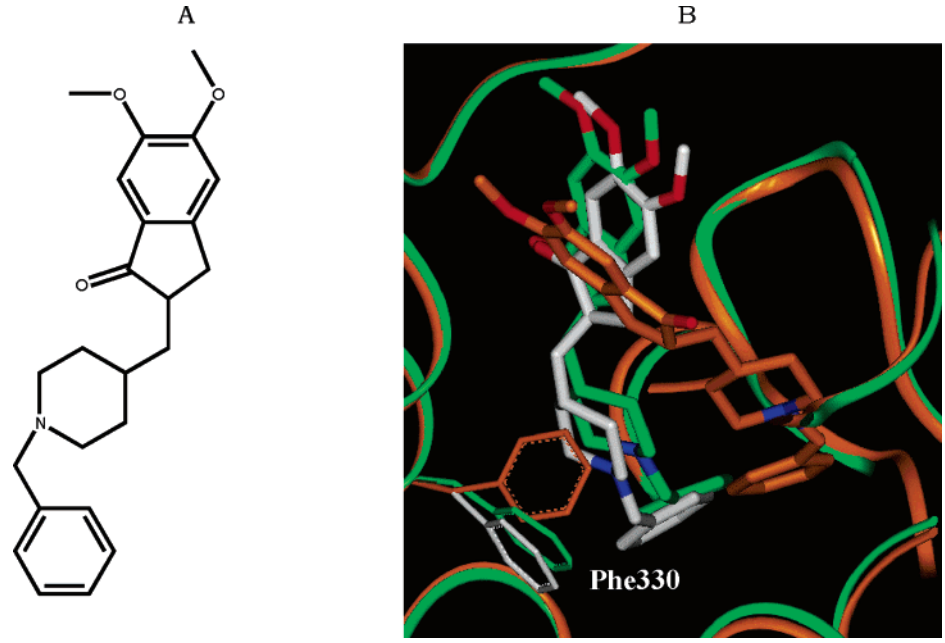

Figure 1. Docking of donepezil to the non-native receptor 1vot: (A) donepezil and (B) the top-ranked docking solution in the rigid and flexible 1vot sites. The crystal structure of the ligand (green) is compared to the top-ranked docking solution in the static 1vot binding site (orange) and the flexible 1vot site (gray). The backbone ribbons of the native receptor 1eve (green) and 1vot (orange) are shown together with the side-chain conformation of the mobile residue Phe330 in 1eve (green), 1vot (orange), and the site from the top-ranked flexible 1vot simulation (gray).

the ligand crystal structure. Analogous runs into the static apo form 1qo9, however, did not yield a clash-free solution within 5.6 Å of the crystal form.

Incorporating receptor flexibility by allowing random χ angle changes (nonrotamer mode) enabled the correct ligand binding mode to be recovered in the native receptor, 1dx4, and the apo form, 1qo9, to within 1.4 Å of the crystal, as the top-ranking solution, as well as the correct orientation of the nine flexible residues. For the apo structure 1qo9, the protein binding site is transformed and resembles that of 1dx4 (Figure 2). The side chains of Tyr71, Tyr370, and also, subtly, Trp83 have rearranged to accommodate the ligand. Interestingly, the best binding affinities from these flexible receptor simulations (−47 kJ/mol) are lower than that achieved in self-docking to the rigid 1dx4 structure (−39 kJ/mol), due to relatively small differences in side-chain and ligand configuration, which elicit enhanced ScreenScore estimates of the protein−ligand interactions.

The same docking simulations to either the native form 1dx4 or the apo form 1qo9 using either the backbone-dependent or independent rotamer libraries were able to recapture the correct ligand binding mode. This form, however, corresponded to the fifth best unique solution (with either library), with a binding energy of −39 kJ/mol. In comparison, the nonrotamer mode yielded the correct configuration as the top ranking solution, with a binding energy of −47 kJ/mol. This is again a consequence of the rotamer conformations available in the libraries used.

These docking experiments provide valuable validation for the methodology. Using rigid receptor snapshots of 1vot and 1qo9, two known highly potent inhibitors of AChE would not be classified as "hits". Introducing side-chain flexibility, however, allows recovery of the correct binding mode in the solution set with the best binding energy.

MMP-1−Flexible Ligand Docking and Structure Generation. The size and shape of the S1′ binding pocket is believed to be the crucial determinant of the ligand binding properties of MMP-1.[66] Figure 3 shows an overlay of two MMP-1−ligand complex crystal structures, which illustrates the striking conformational mobility of certain residues. In previous works, we explored the conformational flexibility of the S1′ pocket in a screening experiment involving attachment of a set of sulfonyl chloride reagents to a common scaffold in the cavity.[28] Here, we extend those earlier studies by conducting ligand generation simulations in the flexible S1′ pocket. First, we validate our approach for this receptor by performing docking runs of ligand RS-104966 (Figure 3) into its cocrystal receptor, 966c, and into the non-native crystal structure, 2tcl. Ligand generation studies are then performed in the S1′ pocket of 2tcl, involving attachment of the "growing" ligand to an anchored hydroxamate molecular scaffold.

Residues Leu181, Arg214, Val215, Ser239, Tyr240, and Phe242 are allowed conformational flexibility in these simulations as they are able to influence the shape and ligand-binding properties of the S1′ pocket. This is the same set of residues that were prescribed flexibility in our earlier works.[28,72] Furthermore, it is known that the segment including residues 238−244 has more mobility than other regions of the structure.[64]

Docking Simulations in the MMP-1 S1′ Pocket. The binding energy data and rmsd comparisons with the ligand crystal structure are shown in Table 5. The binding mode of the synthetic inhibitor RS-104966 was recovered using the rigid native receptor 966c with an rmsd of 0.9 Å to the experimental form and a binding energy of −54.5 kJ/mol. In contrast, analogous docking runs into the rigid 2tcl non-native receptor were unable to yield a solution closer than 4 Å to the crystal as a consequence of clashes with side-chain atoms of the

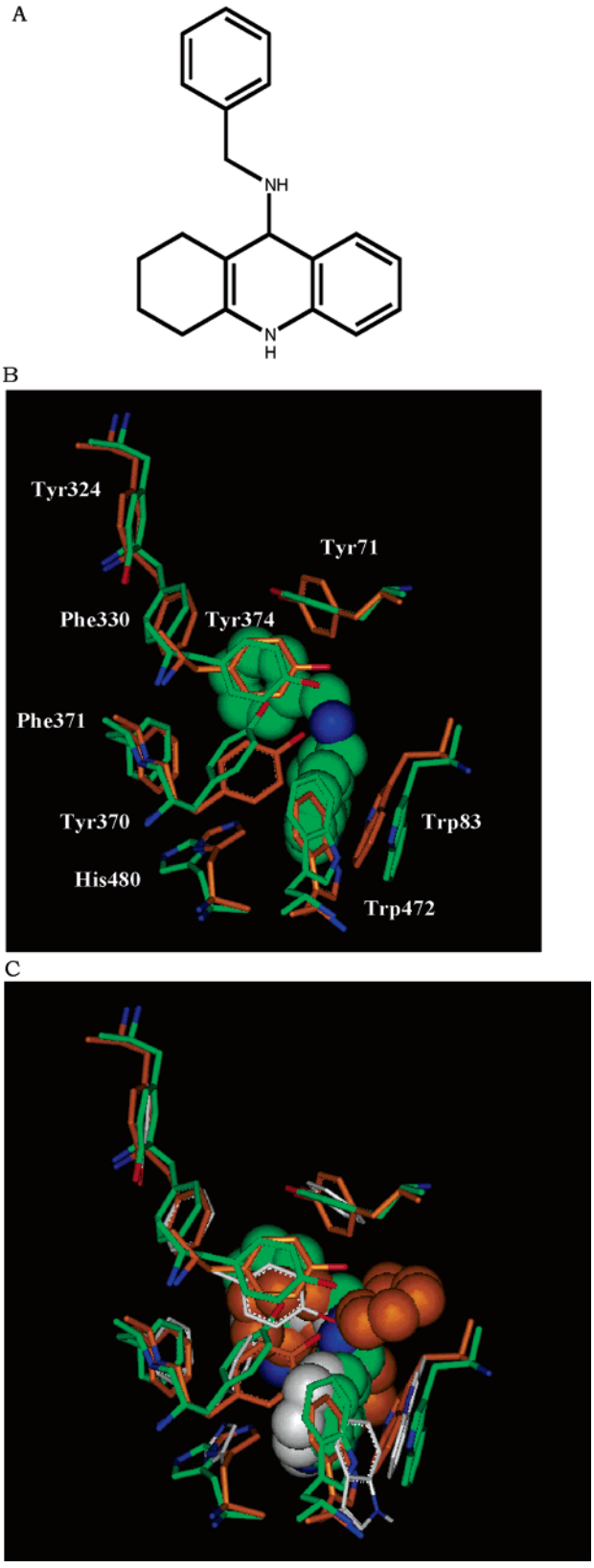

Figure 2. Docking of the tacrine derivative to the apo receptor 1qo9. From left-to-right: (A) tacrine derivative, (B) overlay of the crystal structure of 1dx4 complexed to the tacrine derivative (green) and the apo form 1qo9 (orange), and (C) the top-ranked docking solution in the rigid (orange) and the flexible (gray) 1qo9 sites compared to the ligand crystal structure (green). The side-chain conformations of the nine mobile residues are shown for the cocrystal receptor 1dx4 (green), original 1qo9 form (orange), and the site from the top-ranked flexible 1qo9 simulation (gray).

Table 4. Binding Energies and Binding Mode for Docking of a Tacrine Derivative into Its Cocrystal Receptor 1dx4 and the Apo Form 1qo9

receptor	min energy docking soln/kJ/mol	best rmsd/Å (rank)
rigid 1dx4	−38.6	0.7 (1)
rigid 1qo9[a]	−18.2	5.6 (1)
flexible 1dx4	−47.1	1.3 (1)
flexible 1qo9	−46.9	1.4 (1)

[a] A clash-free solution was not found in the rigid 1qo9 site.

mobile residue Arg114. In fact, the best scoring solution afforded an rms of 4.4 Å with a poor binding energy of −41.9 kJ/mol. In the absence of further structural data, a drug discovery project using a crystal structure snapshot of 2tcl would probably be unable to identify the highly potent RS-104966 inhibitor.

The correct binding mode of RS-104966, to within 1.0 Å of the crystal form, was recovered by docking into the native 966c receptor with the inclusion of side-chain movements in nonrotamer mode. The correct side-chain angles for the six flexible residues were also reproduced. Similarly, by allowing rearrangements of the six side chains in the non-native structure 2tcl, the correct binding mode is recovered to within 1.4 Å of the ligand crystal structure in a modified binding site that resembles the 966c site (Figure 3). The side chain of Arg114 has rotated to open up and to significantly deepen the pocket and allow accommodation of RS-104966 in the experimental binding mode. This arrangement corresponds to the second lowest unique solution in terms of binding energy. The best scoring solution corresponds to a different orientation of RS-104966 in the binding site with an rmsd of 3.5 Å to the experimental conformation. This is similar to the earlier work of Källblad and Dean[72] in which a core ensemble of fifteen different conformations of the S1′ pocket of 2tcl was generated by allowing rotameric variations of the same 6 residue side chains. Docking of RS-104966 into the core ensemble also reproduces the experimental binding mode, but again not as the lowest energy structure. This emphasizes the importance of analyzing a set of low energy binding modes, to increase the probability of including the correct form,[73] and may be a consequence of the quality of the scoring function, which is used to provide an approximate assessment of ligand binding energies.

For this example, docking simulations incorporating side-chain variations through the use of rotamer libraries give very similar results to the analogous runs using random χ angle changes. This is because appropriate rotamers, particularly those for Arg114, that open the S1′ cavity are present in both the backbone-dependent and independent libraries used in this work.

This application validates the approach for allowing receptor flexibility for the docking of RS-104966 into the S1′ cavity of MMP-1. Using the X-ray crystal form of 2tcl leads to a binding mode with a poor energy score that does not match the known bound ligand structure. However, introducing protein structural rearrangements leads to the location of the observed binding mode, which interacts strongly with the S1′ pocket of an alternative receptor conformation.

Structure Generation in the MMP-1 S1′ Pocket. Ligands of different size and shape are known to induce

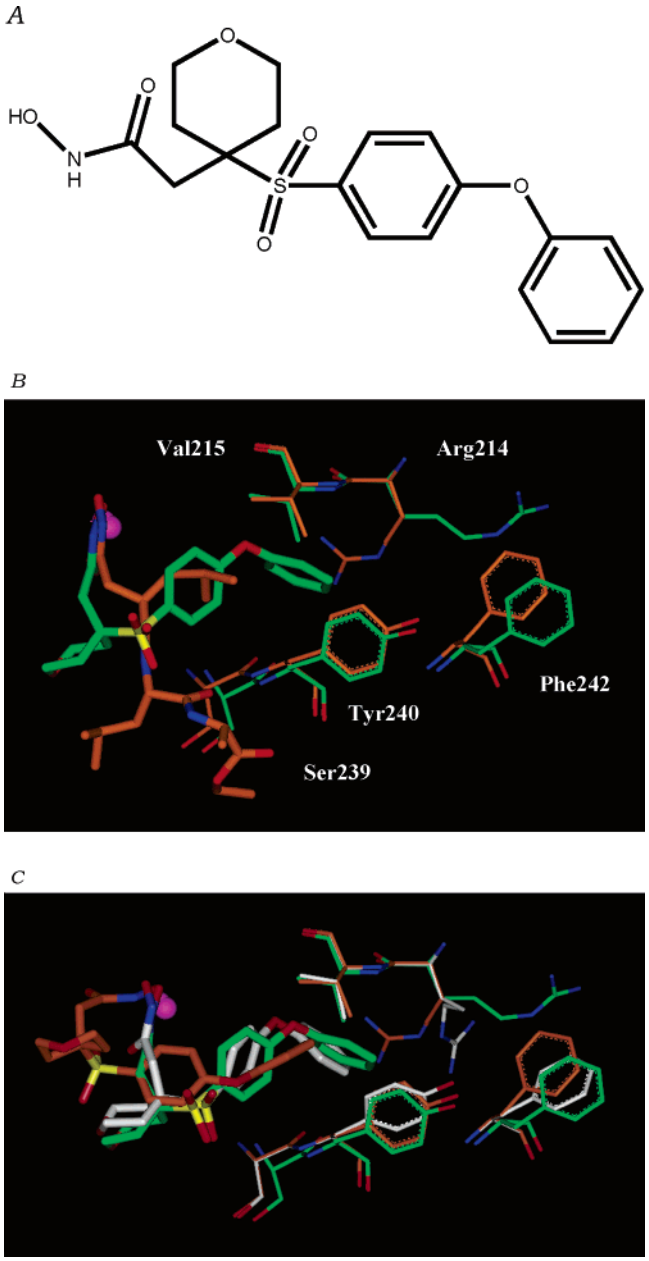

Figure 3. Docking of RS-104966 to the non-native receptor 2tcl: (A) RS-104966, (B) overlay of the crystal structure of 966c complexed to RS-104966 (green) and 2tcl complexed to RO 31-4724 (orange), and (C) the top-ranked docking solution in the rigid (orange) and the flexible (gray) 2tcl sites compared to the ligand crystal structure (green). The side-chain conformations of the mobile residues (except for Leu181) are shown for the cocrystal receptor 966c (green), original 2tcl form (orange), and the site from the top-ranked flexible 2tcl simulation (gray). The zinc atoms are shown in purple.

conformational changes in the S1′ pocket upon binding to MMP-1. We conducted de novo structure generation runs within the S1′ pocket to examine the effect of receptor flexibility on the diversity of the ligands formed. Most of the known MMP inhibitors comprise a functional group, typically a hydroxamate or carboxylate, that forms strong, specific interactions with the catalytic zinc ion. In our simulations, we did not endeavor to regenerate the metal-binding moieties. Instead, a hydroxamate molecular scaffold was rigidly positioned in the site exploiting the coordinates of 2tcl

Table 5. Binding Energies and Binding Mode for Docking of the Synthetic Inhibitor RS-104966 into Its Cocrystal Receptor 966c and the Non-Native Form 2tcl

receptor	min energy docking soln/kJ/mol	best rmsd/Å (rank)
rigid 966c	−51.5	0.9 (1)
rigid 2tcl	−41.9	4.4 (1)
flexible 966c	−53.9	1.0 (1)
flexible 2tcl	−51.6	1.4 (2)

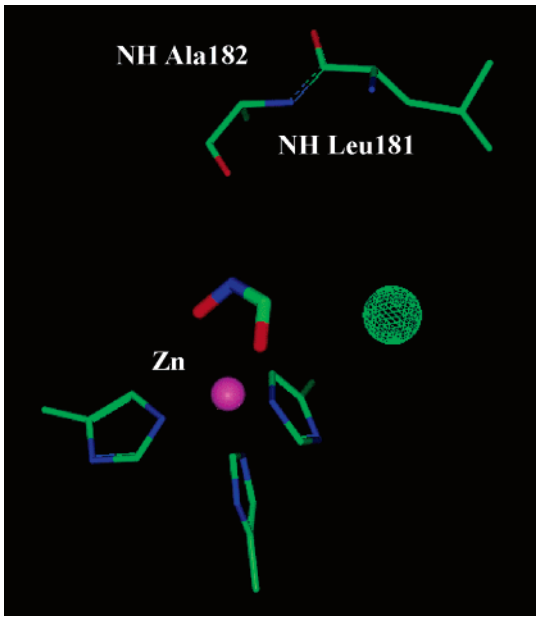

Figure 4. Scaffold and pharmacophore constraints used in the MMP-1 structure generation. The constraints were (1) utilization of a rigid hydroxamate scaffold, (2) hydrogen bond formation to either the Leu181 or Ala182 NH group, and (3) a lipophilic atom within the sphere in the S1′ cavity.

and all structure generation runs involved attachment of the evolving ligand to the scaffold and extension into the S1′ cavity, as depicted in Figure 4. Two pharmacophoric restraints were also supplied as input; a ligand acceptor atom was required to form a hydrogen bond to either the Leu81 or Ala82 backbone NH donor, and a lipophilic atom was required in the S1′ pocket. These constraints are the same as those applied previously to structure generation studies in a rigid MMP-3 site,[38] and they exploit the available knowledge of MMP ligands and their binding characteristics. Ligand design was conducted in both the rigid and flexible 2tcl cavities.

Several known inhibitors for MMP-1 extracted from the WDI are depicted in Figure 5. They generally comprise a carbonyl or sulfonyl group that can form a hydrogen-bonded interaction to the backbone NH of either Leu81 or Ala82, and the number of bonds between the hydroxamate moiety and the carbonyl or sulfonyl group does not vary significantly. A diverse range of hydrophobic substituents are attached to the carbonyl or sulfonyl group, which extend into the S1′ cavity, including single and multiple aromatic rings, fused bi- and tricycles, directly connected rings, and rings bridged by alkyl or ether groups. We showed previously that molecules with these functionalities are expected to participate in the same H-bonds as designated by the pharmacophoric constraints imposed here

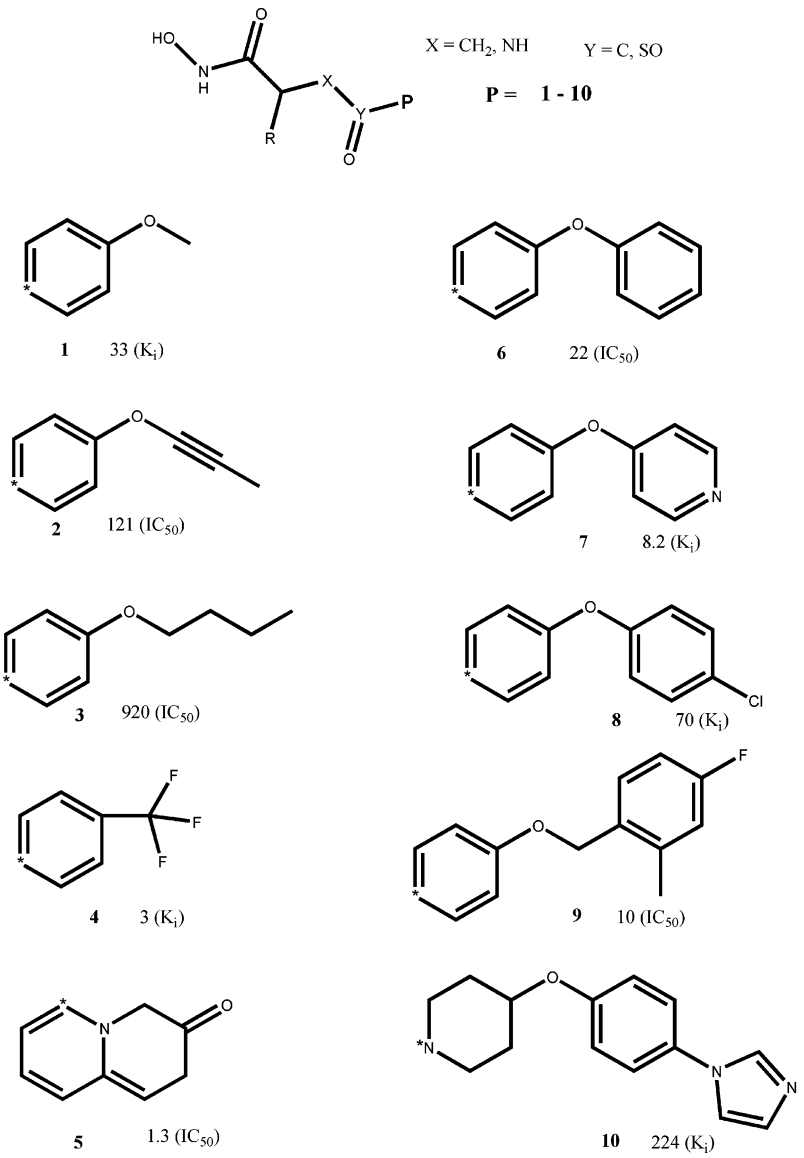

Figure 5. A selection of known MMP-1 inhibitors extracted from the WDI. Activities are given beside each structure as K_i or IC_{50} values in nM. The attachment point to the general scaffold is given by the * symbol.

and their sulfonyl or carbonyl substituents occupy the S1′ cavity.[28] The species attached to the hydroxamate scaffold generated in this work are, thus, expected to have similar binding characteristics to the known inhibitors, and we can directly compare the chemotypes from the rigid and flexible structural design runs with the known molecules.

The designed structures all involve a carbonyl or sulfonyl moiety that satisfies the pharmacophoric constraints with two atom linkers between this group and the hydroxamate. These atomic groups are not directly influenced by the flexible section of the receptor and, thus, similar groups are generated for this portion of the ligands in the static and flexible sites. The difference between the structures generated in the rigid and mobile sites, however, corresponds to differences in the hydrophobic substituents that extend into the S1′ pocket. In the rigid 2tcl receptor, the hydrophobic group in the S1′ cavity typically comprised a single ring with relatively few substituents. Ten of the top scoring ligands are depicted in Figure 6, and their estimated binding energies range from −35 to −47 kJ/mol. The relatively small ligand segments attached to the anchored template can "fit" in the unmodified S1′ pocket configuration.

As discussed above, incorporation of side-chain mobility in the S1′ pocket opens up the cavity. In fact, Källblad et al.[28] quantify the volumes in the S1′ cavity upon receptor conformational changes as ranging from 43 to 148 Å^3. This allows larger structures attached to the hydroxamate anchor and carbonyl/sulfonyl moieties to be formed in the cavity, including biaryl rings, an aliphatic and aromatic ring pair, or fused bi- and tricyclic systems, with a variety of ring connection and substituent groups. Ten of the top-scoring ligands generated are shown in Figure 7, with receptor−ligand binding energies ranging from −45 to −54 kJ/mol.

The molecules generated in the rigid site show functional group similarities with the smaller sets of known MMP-1 inhibitors (1−4 in Figure 5) that were extracted from the WDI. As discussed earlier, the generated and known molecules are deemed to have similar binding characteristics. However, the top-ranking molecules constructed in the flexible site re-

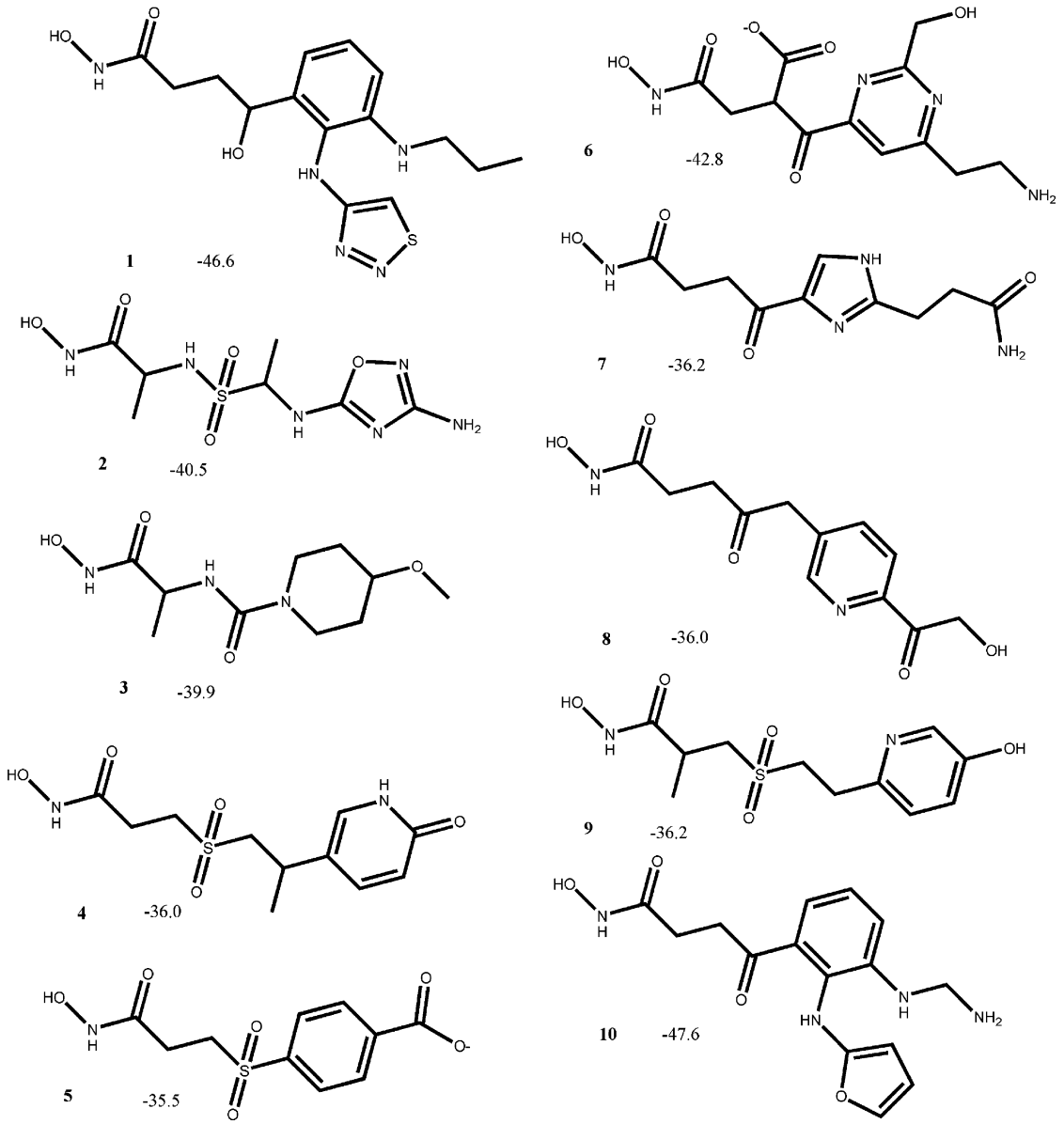

Figure 6. A selection of the top-scoring ligands generated in the rigid 2tcl site. Ligand binding energies are given below each structure.

semble the larger WDI inhibitors (5−10 in Figure 5) that were highly potent and showed more activity, according to the WDI data, than the species resembling the smaller ligands generated in the fixed receptor. This is in agreement with the enhanced binding energies obtained for the larger molecules constructed in the flexible cavity. These molecules do not fit into the rigid S1′ X-ray structure as a result of demanding steric clashes. In most cases, concerted rotation of the Arg114, Tyr140, and Phe142 side chains increases the volume of the cavity and allows larger structures to be accommodated without steric clashes. It should also be noted that the structure generation runs in the flexible 2tcl site also yielded sets of smaller ligands, with lower binding energies, that are similar to those from the runs with the rigid receptor.

For these structure generation experiments, the use of rotamers to represent side-chain conformations gives very similar chemotypes and receptor conformations to the corresponding runs with the random side-chain χ angle (nonrotamer) approach.

The results of these structure generation simulations emphasize the value of allowing for receptor flexibility in a drug-design protocol, particularly if limited protein structural data is available. Ligands of different size and shape, which would score poorly in the original 2tcl structure, can "fit" into a modified S1′ cavity as a result of side-chain rotations. Some of the ligands generated in the flexible receptor approach resemble known highly potent MMP-1 inhibitors that were missed using the rigid receptor. Thus, the utilization of receptor flexibility in a drug discovery project may lead to the development of more active and more diverse ligands than would otherwise be accessible with a static receptor.

Discussion

In this work, we examined methodologies for sampling side-chain conformational space in de novo ligand design and docking. The method integrates receptor conformational sampling into the simulation procedures, and, thus, prior receptor structural explorations are not a necessity. In our approach, ligands are docked or

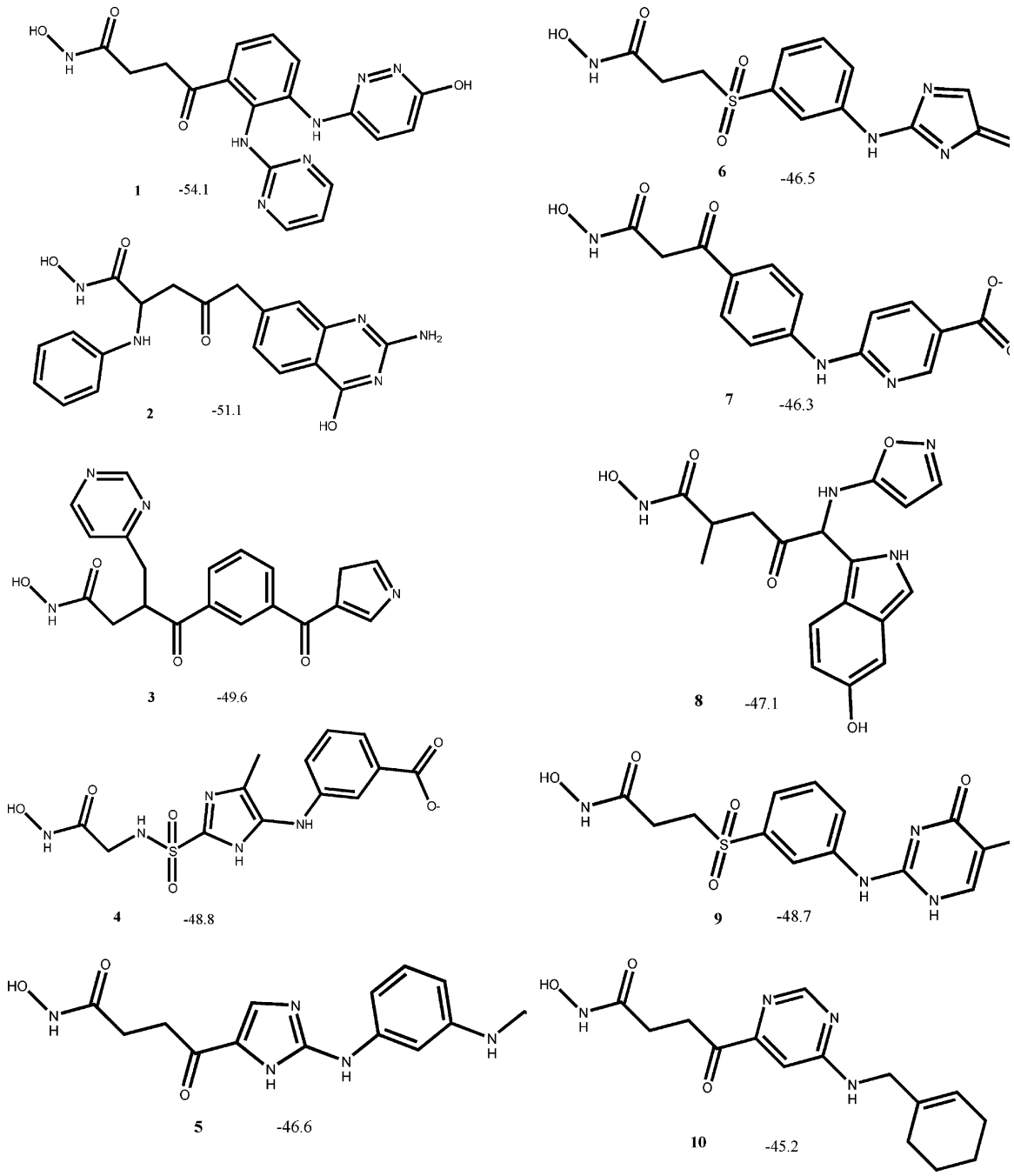

Figure 7. A selection of the top-scoring ligands generated in the flexible 2tcl site. Ligand binding energies are given below each structure.

generated within the target active site using a simulated annealing method to optimize the binding free energy.[38] Local side-chain flexibility was incorporated by allowing changes in rotameric states (from libraries of side-chain rotamer conformations) or variations in side-chain χ dihedral angles (nonrotamer approach). At each Monte Carlo step of the annealing method, if a protein structural transition is chosen, a side chain from among the flexible set is selected and a set of side-chain χ angles or a rotamer representation is randomly assigned to that residue. The evolving ligand is assessed against the new protein structural model, and the transition is accepted or rejected according to the Metropolis condition.

The method for handling side-chain mobility was validated in several ligand docking experiments. It was found that docking of particular AChE and MMP-1 ligands into non-native receptor binding sites that were fixed in their original conformations were unable to reproduce the experimentally known ligand binding modes. In contrast, exploiting our approach for receptor flexibility allowed specific side chains to rotate sufficiently to accommodate these ligands in their crystal structure orientations.

Use of side-chain χ dihedral angle changes to represent protein flexibility proved to be generally successful in these ligand docking and design studies. The alternative approach, using rotamers from predefined rotamer libraries to portray different side-chain conformers, suffers the limitation that appropriate rotameric states have to be present within the library. Furthermore, many different libraries are available containing dif-

ferent rotameric forms, and the best choice of library is not always clear prior to their usage. For a particular project, perhaps multiple libraries should be evaluated. In our studies, we used the same Monte Carlo length for both rotamer and random χ angle change simulations. From the results, which, overall, were more successful for the χ angle change representation of protein flexibility, it is apparent that the latter approach did not suffer from conformational sampling problems. For these reasons, it is probably safest, convenient, and more efficient to use the random χ angle nonrotamer mode for incorporating receptor flexibility.

Ligand generation studies were conducted within a rigid and flexible MMP-1 S1′ cavity, in which the evolving ligands were attached to an anchored hydroxamate molecular template and were constrained to satisfy certain pharmacophoric constraints. The fragment-based design method used in this study leads to a body of molecules that show similarities with known active MMP-1 inhibitors. Use of the rigid 2tcl receptor leads to the production of relatively small hydrophobic groups in the S1′ cavity, typically comprising a single ring with relatively few substituents. However, incorporation of receptor flexibility opened up the cavity and allowed significantly larger structural moieties with better binding energies to be "grown" in that region of the binding site, typically comprising two or more rings with different ring connections and substituents. The flexible receptor approach also leads to strongly binding molecules, resembling some of the potent inhibitors that were found to be "hits" in our earlier reagent screening project,[28] that do not "fit" into the static site.

Protein conformational flexibility has recently been introduced into docking procedures. Most previous de novo structure-based design methods, however, consider only a single, rigid protein structure, and, as a result of the huge conformational search space, the question of protein structural flexibility is generally not addressed. The methodology discussed in this work for handling receptor flexibility has been shown to be valuable in both de novo ligand design and docking scenarios. It provides a fast, efficient, and valid approach for the integration of protein conformational rearrangements into the drug discovery process to enhance diversity coverage and activity of the designed ligands. Of course, its weakness is that it focuses on side-chain mobility and neglects backbone movements, which not only influence side-chain conformations but can significantly change the shape of the binding cavity. Nevertheless, it is well-known that side-chain orientations are less conserved than backbone configurations and, thus, the incorporation of side-chain mobility represents an important step toward the handling of fully flexible protein structures.

References

(1) Jorgensen, W. L. The many roles of computation in drug discovery. *Science* **2004**, *303*, 1813−1818.

(2) Teague, S. J. Implications of protein flexibility for drug discovery. *Nat. Rev. Drug Discovery* **2003**, *2*, 527−540.

(3) Vieira, E.; Binggeli, A.; Breu, V.; Bur, D.; Fischli, W.; Guller, R.; Hirth G.; Marki, H. P.; Muller, M.; Oefner, C.; Scalone, M.; Stadler, H.; Wilhelm, M.; Wostl, W. Substituted piperidines−highly potent renin inhibitors due to induced fit adaptation of the active site. *Bioorg. Med. Chem. Lett.* **1999**, *9*, 1397−1402.

(4) Weichsel, A.; Montfrot, W. Ligand-induced distortion of an active site in thymidylate synthase upon binding anticancer drug 1843U89. *Nat. Struct. Biol.* **1995**, *2*, 1095−1101.

(5) Tong, L.; Pav, S.; Suet, M.; Lamarre, D.; Yoakim, C.; Beaulieu, P.; Anderson, P. Crystal structures of HIV-2 protease in complex with inhibitors containing the hydroxyethylene dipeptide isostere. *Structure* **1995**, *3*, 33−40.

(6) Najmanovich, R.; Kuttner, J.; Sobolev, V.; Edelman, M. Side-chain flexibility in proteins upon ligand binding. *Proteins: Struct., Funct., Genet*. **2000**, *39*, 261−268.

(7) Lesk, A. M.; Chothia, C. Elbow motion in immunoglobins involves a molecular ball-and-socket joint. *Nature* **1988**, *335*, 188−190.

(8) Ewing, T. J. A.; Makino, S.; Skillman, A. G.; Kuntz, I. D. DOCK 4.0; Search strategies for automated molecular docking of flexible molecule databases. *J. Comput.-Aided Mol. Des.* **2001**, *15*, 411−428.

(9) Morris, G. M.; Goodsell, D. S.; Halliday, R. S.; Huey, R.; Hart, W. E.; Betlew, R. K.; Olson, A. J. Automated docking using a Lamarckian genetic algorithm and an empirical binding free energy function. *J. Comput. Chem.* **1998**, *19*, 1639−1662.

(10) Rarey, M.; Kramer, B.; Lengauer, T.; Klebe, G. A fast flexible docking method using an incremental construction algorithm. *J. Mol. Biol*. **1996**, *261*, 470−489.

(11) Jones, G.; Willet, P.; Glen, R. C.; Leach, A. R.; Taylor, R. Development and validation of a genetic algorithm for flexible docking. *J. Mol. Biol*. **1997**, *267*, 727−748.

(12) Friesner, R. A.; Banks, J. L.; Murphy, R. B.; Halgren, T. A.; Klicic, J. J.; Mainz, D. T.; Repasky, M. P.; Knoll, E. H.; Shelley, M.; Perry, J. K.; Shaw, D. E.; Francis, P.; Shenkin, P. S. Glide: A new approach for rapid, accurate docking and scoring. 1. Method and assessment of docking accuracy. *J. Med. Chem*. **2004**, *47*, 1739−1749.

(13) Carlson, H. A.; McCammon, J. A. Accommodating protein flexibility in computational drug design. *Mol. Pharmacol.* **2000**, *57*, 213−218.

(14) Yoon, S.; Welsh, W. J. Identification of a minimal subset of receptor conformations for improved multiple conformation docking and two-step scoring. *J. Chem. Inf. Comput. Sci*. **2004**, 44, 88−96.

(15) Apostolakis, J.; Pluckthun, A.; Caflisch, A. Docking Small Ligands in Flexible Binding-sites. *J. Comput. Chem.* **1998**, *19*, 21−37.

(16) Cavasotto, C. N.; Abagyan, R. A. Protein flexibility in ligand docking and virtual screening to protein kinases. *J. Mol. Biol*. **2004**, *337*, 209−225.

(17) Ragno, R.; Frasca, S.; Manetti, F.; Brizzi A.; Massa, S. HIV-Reverse transcriptase inhibition: inclusion of ligand-induced fit by cross-docking studies. *J. Med. Chem*. **2004**.

(18) Vigers, G. P. A.; Rizzi, J. P. Multiple active site corrections for docking and virtual screening. *J. Med. Chem*. **2004**, *47*, 80−89.

(19) Fernandez, M. X.; Kairys, V.; Gilson, M. K. Comparing ligand interactions with multiple receptors via serial docking. *J. Chem. Inf. Comput. Sci*. **2004**, *44*, 1961−1970.

(20) Knegtel, R. M. A.; Kuntz, I. D.; Oshiro, C. M. Molecular docking to ensembles of protein structures. *J. Mol. Biol.* **1997**, *266*, 424−440.

(21) Österberg, F.; Morris, G. M.; Sanner, M. F.; Olsen, A. J.; Goodsell, D. S. Automated docking to multiple targets structures: incorporation of protein mobility and structural water heterogeneity in AutoDock. *Proteins: Struct., Funct., Genet.* **2002**, *46*, 34−40.

(22) Broughton, H. B. A method for including protein flexibility in protein−ligand docking: improving tools for database mining and virtual screening. *J. Mol. Graphics Modell*. **2000**, *18*, 247−257.

(23) Totrov, M.; Abagyan, R. Flexible protein-ligand docking by global energy optimization in internal co-ordinates. *Proteins: Struct., Funct., Genet.* **1997**, *29*, 215−220.

(24) Schnecke, V.; Kuhn, L. A. Virtual screening with solvation and ligand-induced complementarity. *Perspect. Drug Discov*. **2000**, *20*, 171−190.

(25) Leach, A. R. Ligand docking to proteins with discrete side-chain flexibility. *J. Mol. Biol.* **1994**, *235*, 345−356.

(26) Frimurer, T. M.; Peters, G. H.; Iversen, L. F.; Andersen, H. S.; Moller, N. P.; Olsen, O. H. Ligand-induced conformational changes: improved predictions of ligand binding conformations and affinities. *Biophys. J.* **2003**, *84*, 2273−2281.

(27) Taylor, R. D.; Jewsbury, P. J.; Essex, J. W. FDS: flexible ligand and receptor docking with a continuum solvent and soft core energy function. *J. Comput. Chem*. **2003**, *24*, 1637−1656.

(28) Källblad, P.; Todorov, N. P.; Willems, H. M. G.; Alberts, I. L. Receptor flexibility in the in silico screening of reagents in the S1′ pocket of human collagenase. *J. Med. Chem*. **2004**, *47*, 2761−2767.

(29) Lill, M. A.; Vedani, A.; Dobler, M. Raptor: Combining dual-shell representation, induced fit simulation, and hydrophobicity scoring in receptor modeling: Application toward the simulation of structurally diverse ligand sets. *J. Med. Chem*. **2004**, *47*, 6174−6186.

(30) Bouzida, D.; Rejto, P. A.; Arthurs, S.; Colson, A. B.; Freer, S. T.; Gehlhaar, D. K.; Larson, V.; Luty, B. A.; Rose, P. W.; Verkhivker, G. M. Computer simulations of ligand-protein binding with ensembles of protein conformations: A Monte Carlo study of HIV-1 protease binding energy landscapes. *Int. J. Quantum Chem.* **1999**, *72*, 73−84.
(31) Murray, C. W.; Baxter, C. A.; Frenkel, A. D. The sensitivity of the results of molecular docking to induced fit effects: application to thrombin, thermolysin and neuraminidase. *J. Comput.-Aided Mol. Des.* **1999**, *13*, 547−562.
(32) Carlson, H. A. Protein flexibility is an important component of structure-based drug discovery. *Curr. Pharm. Dis*. **2002**, *8*, 1571−1578.
(33) Bursavich, M. G.; Rich, E. H. Designing non-peptide peptidomimetics in the 21st century: inhibitors targeting conformational ensembles. *J. Med. Chem*. **2002**, *45*, 541−558.
(34) Zhu, J.; Fan, H.; Liu, H.; Shi, Y. Structure-based ligand design for flexible proteins: Application of new F-DycoBlock. *J. Comput.-Aided Mol. Des*. **2001**, *15*, 979−996.
(35) Böhm, H.-J. The computer program LUDI: a new method for the de novo design of enzyme inhibitors. *J. Comput.-Aided Mol. Des.* **1992**, *6*, 61−78.
(36) Eisen, M. B.; Wiley, D. C.; Karplus, M.; Hubbard, R. E. HOOK: a program for finding novel molecular architectures that satisfy the chemical and steric requirements of a macromolecule binding-site. *Proteins: Struct., Funct., Genet.* **1994**, *19*, 199−221.
(37) Gillet, V.; Johnson, A. P.; Mata, P.; Sike, S.; Williams, P. SPROUT: a program for structure generation. *J. Comput.-Aided. Mol. Des.* **1994**, *7*, 127−153.
(38) Stahl, M.; Todorov, N. P.; James, T.; Mauser, H.; Boehm, H. J.; Dean, P. M. A validation study of the practical use of automated de novo design. *J. Comput.-Aided Mol. Des.* **2002**, *16*, 459−478.
(39) Chung, S. Y.; Subbiah, S. How similar must a template protein be for homology modelling by side-chain packing methods? In *Pacific Symposium On Biocomputing*; Hunter, L., Klein, T. E., Eds.; World Scientific: Hawaii, 1996; pp 126−141.
(40) Najmanovich, R.; Kuttner, J.; Sobolev, V.; Edelman, M. Side-chain flexibility in proteins upon ligand binding. *Proteins: Struct., Funct., Genet*. **2000**, *39*, 261−268.
(41) Ponder, J. W.; Richards, F. M. Tertiary templates for proteins—use of packing criteria in the enumeration of allowed sequences for different structural classes. *J. Mol. Biol*. **1987**, *193*, 775−791.
(42) Dunbrack, R. L; Karplus, M. Back-bone-dependent rotamer library for proteins—application to side-chain prediction. *J. Mol. Biol*. **1993**, *230*, 543−574.
(43) Dunbrack, R. L; Cohen, F. E. Bayesian statistical analysis of protein side-chain rotamer preferences. *Protein Sci.* **1997**, *6*, 1661−1681.
(44) Xiang, Z.; Honig, B. Extending the Accuracy Limits of Prediction for Side-chain Conformations. *J. Mol. Biol*. **2001**, *311*, 421−430.
(45) Lovell, S. C.; Word, J. M.; Richardson, J. S.; Richardson D. C. The penultimate rotamer library. *Proteins: Struct., Funct., Genet*. **2000**, *40*, 389−408.
(46) Sali, A.; Blundell, T. L. Comparative protein modeling by satisfaction of spatial constraints. *J. Mol. Biol*. **1993**, *234*, 779−815.
(47) Bower, M.; Cohen, F.; Dunbrack, R. J. Prediction of protein side-chain rotamers from a back-bone-dependent rotamer library: a new homology modeling tool. *J. Mol. Biol*. **1997**, *267*, 1268−1282.
(48) Dahiyat, B. I; Mayo, S. L. De novo protein design: fully automated sequence selection. *Science* **1997**, *278*, 82−87.
(49) Wernisch, L.; Hery, S.; Wodak, S. J. Automatic protein design with all atom force-fields by exact and heuristic optimization. *J. Mol. Biol*. **2000**, *301*, 713−736.
(50) Desjarlais, J. R.; Handel, T. M. Side-chain and back-bone flexibility in protein core design. *J. Mol. Biol*. **1999**, *290*, 305−318.
(51) Lee, C.; Subbiah, S. Prediction of protein side-chain conformation by packing optimization. *J. Mol. Biol*. **1991**, *217*, 272−288.
(52) Desmet, J.; De Maeyer, M.; Hazes, B.; Lasters, I. The dead-end elimination theorem and its use in protein side-chain positioning. *Nature* **1992**, *356*, 539−542.
(53) Lasters, I.; Desmet, J. The fuzzy-end elimination theorem: correctly implementing the side-chain placement algorithm based on the dead-end elimination theorem. *Protein Eng*. **1993**, *6*, 717−722.
(54) Keller, D.; Shibata, M.; Markus, E.; Ornstein, R.; Rein, R. Finding the global minimum: a fuzzy end elimination implementation. *Protein Eng*. **1995**, *8*, 893−904.
(55) Todorov, N. P.; Dean, P. M. An evaluation of a method for controlling molecular scaffold diversity in *de novo* ligand design. *J. Comput.-Aided Mol. Des.* **1997**, *11*, 175−192.
(56) Todorov, N. P.; Dean, P. M. Computational Ligand Design by Free Energy Minimization. In *Biological Physics*; Frauenfelder, H., Hummer, G., Garcia, R., Eds.; American Institute of Physics: New York, 1999.
(57) Kryger, G.; Silman, I.; Sussman, J. L. Structure of acetylcholinesterase complexed with E2020 (Donepezil): implications for the design of new anti-Alzheimer drugs. *Structure* **1999**, *7*, 297−307.
(58) Raves, M. L.; Harel, M.; Pang, Y.-P.; Silman, I.; Kozikowski, A. P.; Sussman, J. L. Structure of acetylcholinesterase complexed with the nootropic alkaloid, (−)-huperzine A. *Nat. Struct. Biol.* **1997**, *4*, 57−63.
(59) Harel, M.; Kryger, G.; Rosenberry, T. R.; Mallender, W. D.; Lewis, T.; Fletcher, R. J.; Guss, J. M.; Silman, I.; Sussman, J. L. Three-dimensionsal structures of Drosophila melanogaster acetylcholinesterase and of its complexes with two potent inhibitors. *Protein Sci*. **2000**, *9*, 1063−1072.
(60) Coussence, L. M.; Fingleton, B.; Matrisian, L. M. Matrix metalloproteinase inhibitors and cancer: trials and tribulations. *Science* **2002**, *295*, 2387−2392.
(61) Cleutjens, J. P. The role of matrix metalloproteinases in heart disease. *Cardiovasc. Res*. **1996**, *30*, 816−821.
(62) Beckett, R. P.; Whittaker, M. Matrix metalloproteinase inhibitors. *Expert Opin. Ther. Pat*. **1998**, *8*, 259−282.
(63) Babine, R. E.; Bender, S. L. Molecular recognition of protein-ligand complexes: Applications to drug design. *Chem. Rev*. **1997**, *97*, 1359−1472.
(64) Moy, F. J.; Chanda, P. K.; Chen, J. M.; Cosmi, S.; Edris, W.; Skotnicki, J. W.; Powers, R. Solution structure of the catalytic fragment of human fibroblast collagenase complexed with a sulphonamide derivative of a hydroxamic acid compound. *Biochemistry* **1999**, *38*, 7085−7096.
(65) Botos, I.; Scapozza, L.; Zhang, D.; Liotta, L. A.; Meyer, E. F. Batimastat, a potent matrix metalloproteinase inhibitor, exhibits an unexpected mode of binding. *Proc. Natl. Acad. Sci. U.S.A.* **1996**, *93*, 2749−2754.
(66) Lovejoy, B.; Welsh, A. R.; Carr, S.; Luong, C.; Broka, C.; Hendricks, R. T. Campbell, J. A.; Walker, K. A. M. Martin, R.; van Wart, H.; et al. Crystal structures of MMP-1 and -13 reveal the structural basis for selectivity of collagenase inhibitors. *Nat. Struct. Biol*. **1999**, *6*, 217−221.
(67) Maple, J. R.; Dinur, U.; Hegler, E. T. Derivation of force fields for molecular mechanics and dynamics from ab initio energy surfaces. *Proc. Natl. Acad. Sci. U.S.A.* **1988**, *85*, 5350−5354.
(68) Bondi, A. van der Waals volumes and radii. *J. Phys. Chem*. **1968**, *68*, 441−451.
(69) Stahl, M.; Rarey, M. Detailed Analysis of Scoring Functions for Virtual Screening. *J. Med. Chem.* **2001**, *44*, 1035−1042.
(70) Lewell, X. Q.; Judd, D. B.; Watson, S. P.; Hann, M. M. RECAP: retrosynthetic combinatorial analysis procedure: a powerful new technique for identifying privileged molecular fragments with useful applications in combinatorial chemistry. *J. Chem. Inf. Comput. Sci.* **1998**, *38*, 511−522.
(71) Erickson, J. A.; Jalaie, M.; Robertson, D. H.; Lewis, R. A.; Vieth, M. Lessons in molecular recognition: The effects of ligand and protein flexibility on molecular docking accuracy. *J. Med. Chem.* **2004**, *47*, 45−55.
(72) Källblad, P.; Dean, P. M. Efficient sampling of local side-chain flexibility. *J. Mol. Biol*. **2003**, *326*, 1651−1665.
(73) Källblad, P.; Mancera, R. L.; Todorov, N. Assessment of Ligand Binding Modes in Protein-Ligand Docking. *J. Med. Chem.* **2004**, *47*, 3334−3337.
(74) Borkakoti, N.; Winkler, F. K.; Williams, D. H.; Darcy, A.; Broadhurst, M. J.; Brown, P. A.; Johnson, W. H.; Murray, E. J. Structure of the catalytic domain of human fibroblast collagenase complexed with an inhibitor. *Nat. Struct. Biol*. **1994**, *1*, 106−110.

JM050196J